Comparison of LMS and NLMS algorithm with the using of 4 Linear Microphone Array for Speech Enhancement

Mamun Ahmed, Nasimul Hyder Maruf Bhuyan

Abstract— In this paper, we have presented the design, implementation and comparison result of Least Mean Square (LMS) algorithm and Normalized LMS (NLMS) algorithm using a 4 channel microphone array for noise reduction as well as speech enhancement. Adaptive sub band Generalized Side lobe Canceller (GSC) beam former has been used for experiment and analysis. Tested results were done by using one speech signal and a small number of noise sources. The side lobe canceller was evaluated with the adaptation of LMS and NLMS. The overall development of Signal to Noise Ratio (SNR) has been determined from the input and output powers of signal and noise, with signal only as input and noise, as input to the GSC. The NLMS algorithm considerably improves speech quality with noise suppression levels of up to 13 dB, while the LMS algorithm is giving up to 10 dB. In different ways of SNR measure was under various types of blocking matrix, step sizes and various noise locations. The whole process will be used for hands free telephony, video conferencing etc. in a noisy environment.

Index Terms— Least Mean Square (LMS); Normalized LMS (NLMS); Generalized Side lobe Canceller (GSC); SNR; Microphone array.

I. INTRODUCTION

Some difficult voice applications, like as video conferencing and sensible rooms, can use microphones which will be several meters away from the speakers. Under these conditions, recorded signals are harshly degraded by noise and echo, and a few kinds of processing is sometimes necessary to improve the speech signal [1]. Adaptive beam formation based on a GSC has been widely thought of due to its effectiveness and ease in accomplishing multiple shaped linearly restricted and to some extent adaptable forms. Nevertheless, many reports show that adaptive beam formers based on GSC are often very sensitive to steering angle misalignments and weight vectors [2].

The main contributions of this paper are comparable of LMS and Normalized LMS algorithm for noise reduction as well as speech enhancement with the help of sub band GSC beam forming.

II. THEORETICAL CONSIDERATIONS

A. Fundamentals of sub band

If a desire system state is labeled as $\{X_k; k = 0, 1,. . \}$, Where $X_k \in \{0, 1\}_d$ is a Boolean vector with dimension d - the representation network of the discrete time k, activation / inactivation state of genes containing X_k components in the present case of a genes regulator network.

The situation is assumed to be updated and witnessed separately. Non-linear time signal through the following model:

$$X_k = f (X_{k-1}) \oplus n_k \text{ (state model)} \tag{1}$$
$$Y_K = h (X_k, v_k) \text{ (observation model)} \tag{2}$$

Where the Y_k is a vector and observation noise v_k, which is using for measurement. "$\oplus$" specifies the component wise modulation-2. Noise processes $\{n_k v_k; k = 1, 2,. . .\}$. It is assumed to be "white" in the sense which is uncorrelated each other [3].

We used additive white Gaussian noise like as random and it was mixed with different SNR (5dB, 10dB, 12dB) for measure the overall SNR development.

When the entire band signal is split into sub bands, sample it at a lower rate because of the reduced bandwidth. The resulting separate sub-bands may be individually processed in the post-processing, like as audio coding. Using the synthesis filter bank to a full band system output at the base sampling rate these sub band signal processing can be reconstructed. Different sampling rates are often used in different parts of the system; they are called as multi-rate filter bank.

There are 2 main advantages when using sub band adaptive beam forming. First one is the convergence speed which was happening because of the pre-whitening result of the sub-band decomposition and another one is in reduced computational complexity due to a lower sampling rate at the decimated sub-bands.

Fig. 1 shows the combinations of a filter bank of *K*-channel with a decimation factor of *N,* where the input signal *x[n]* is decomposed into *K* sub bands by a test filter $H_0(z)… H_{K-1}(z)$ with each sub band down sampled by a factor of $N \le K$ [4].
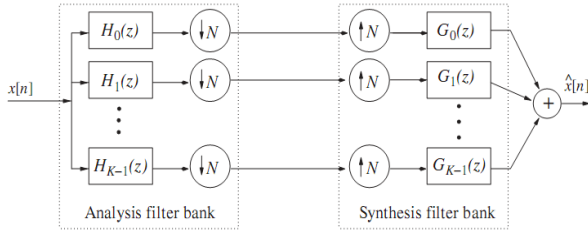
Fig.1. The arrangement of a *K*-channel filter bank with a decimation factor *N*.

B. Problem Formulation and Algorithm

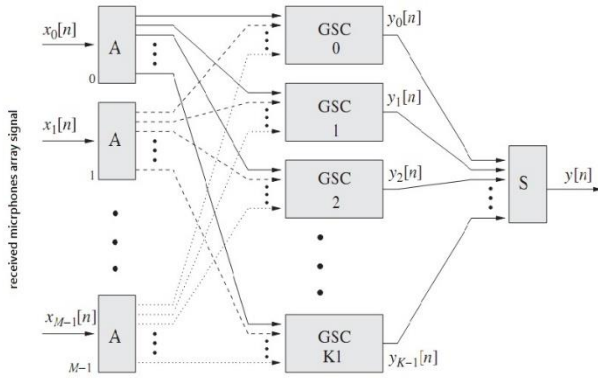Fig. 2 summarizes the overview of the whole system process.

Fig.2. A general Sub band Adaptive Beam forming (SAB) structure with a GSC at each set of the sub bands signal.

Fig. 2 shows a general SAB structure in which each of the received *M* signals $x_m[n]$, *m = 0, 1, ..., M -1* is decomposed into *K* sub bands by a filter bank of *K*-channel analysis and a GSC beam former is then configured in each set of the corresponding decimated *M*-sub band signals. The output y_k *[n]*, *k = 0, 1... K −1*, of these sub band beam formers *K* are then collective by a synthesis filter bank to form the full band output *y[n]*.

Here, label '*A*' are the analysis filter banks (with the down sampling) and the block labeled '*S*' are the synthesis filter bank [5].

Our main target is to extract better SNR as well as speech enhancement in a noisy environment with using of LMS and NLMS algorithm where from a signal source is emitting a voice signal which was situated at in front and middle of microphones and noise was taken from right, left or other positions.
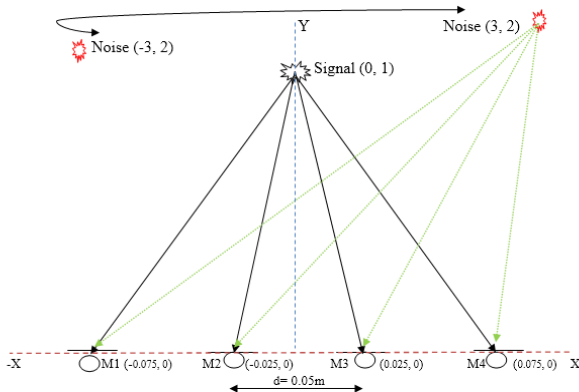

Fig.3. Microphones, Signal and noise position.

C. Formulation of the LMS and NLMS algorithm

LMS algorithm: LMS algorithm's equations can be summarized in the following ways;

Output, $y(n) = w^H x(n)$ (3)
Error signal, $e(n) = d(n) - w^H x(n)$ (4)

d(n) = desired signal; *x(n)* = input signal
Weight update equation,

$$w(n+1) = w(n) + \mu x(n)e^*(n) \tag{5}$$

Where, μ is the step size parameter which controls the convergence properties of the LMS algorithm [6].

NLMS algorithm: The Normalized LMS (NLMS) introduces a variable adaptation rate which improves the convergence speed in a non-static environment [6].

$$W_k(n+1) = W_k(n) + \frac{\alpha e*(n)x(n)}{\|x(n)\|^2} \tag{6}$$

(Where "*" represents the conjugate value and "α" is the normalized step size with $0 < \alpha < 2$)

III. GSC AND OVERVIEW OF THE FULL SYSTEM

The Generalized Side lobe Canceller (GSC) is a simplified type of the Frost algorithm, presented by Griffiths and Jim about ten years after the first Frost document was published [6]. They proposed an alternate but effective implementation of the LCMV beam former, which is called the CGC. The GSC will be thought of as a system to transform the restricted minimization problem into an unrestricted one [4].

In Fig. 4, the structure consists of two parts called the top and bottom part. At the top components is usually known as the fixed beam former and the bottom consists of adaptive section together with the blocking matrix [8].
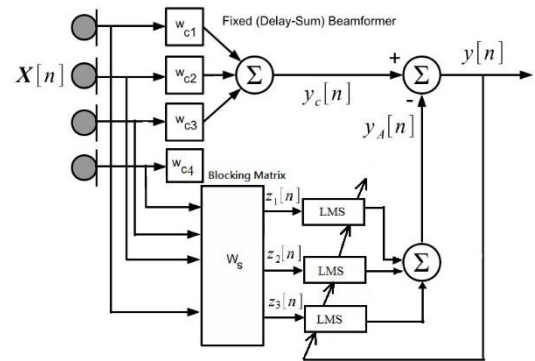

Fig.4. Structure of GSC.

X[n] can be extracted from *M*-1 linearly independent components such as microphones. The standard Griffiths-Jim Blocking Matrix (BM) is [7].

$$Ws = \begin{bmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 \end{bmatrix}$$

For these Ws, the BM outputs are calculated as the matrix product of the BM and the current input matrix data.

$$Z[n] = W_s^H X[n] \qquad (7)$$

The whole beam former output, $y[n]$, is computed as the following way:

$$y[n] = y_c[n] - \sum_{k=1}^{M-1} \boldsymbol{W}_k^H[n] z_k[n] \qquad (8)$$

Where $W_k[n]$ is the kth column of the tap weight matrix W and $z_k[n]$ is the kth blocking matrix output and they have the same length matrix. In the adaptive filters weight updated using the LMS algorithm with reference signal as $y[n]$

$$W_k[n+1] = W_k[n] + \mu y^*[n] z_k[n] \qquad (9)$$

(Where "*" represents the conjugate value)

When we used NLMS algorithm in the GSC, weight update equation is like as below:

$$W_k[n+1] = W_k[n] + \frac{\alpha y^*[n] z_k[n]}{\|z_k[n]\|^2} \qquad (10)$$

In the form of matrix, the blocking matrix formulated as below:
$$\tilde{B} = B_M . B_{M-1} \dots B_{M-S+1} \qquad (11)$$

Where we have

$$B_i = \begin{bmatrix} 1 & -1 & & & 0 \\ & & \cdot & \cdot & \\ & & & \cdot & \cdot \\ 0 & & & 1 & -1 \end{bmatrix}^H$$

With $i = M, M-1 \dots M-S-1$. If the signal of interest comes from the wide side, it will not be possible to undergo such an above blocking matrix. The zero response shaped by the side block matrix will have a wider lobe width with the increase of S [4].

The general sub band adaptive filter system (SAF) as shown in Fig. 5, where the input signal and the desired signal are divided into sub-bands decimated by analysis filter banks and then with the adaptive filters of Sub band as Windowed Discrete Fourier Transform (WDFT), which runs at a much lower rate compared to the original full-band system used to estimate the required sub-band signals using the sub band input signals. The resulting sub band error signals are reconstructed into a full band error signal by a bank of synthesis filters [4].
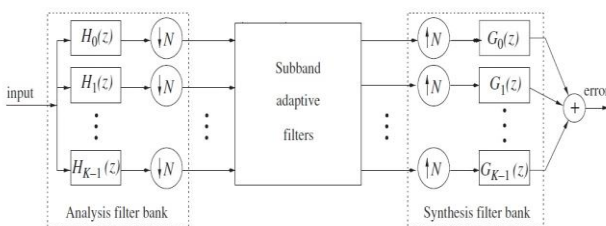

Fig.5. General SAF structure [3].

The WDFT of $x[n]$ can be achieved as [9]

$$X[m, w_i] = \sum_{n=0}^{N-1} S[n] x[Mm+n] e^{-jw_i n} \qquad (12)$$

Where, $w_i = \frac{2\pi i}{N}$, $i = 0 \dots N-1$, and $S[n]$ is a window function. In the framework of filter banks, $X[m, w_i]$ can be regarded as the i:th sub band signal.

The inverse WDFT is usually given as

$$x[Mm+n] = \frac{1}{KN} \sum_{i=0}^{N-1} \sum_{k=0}^{K-1} \frac{1}{S[i]} X[m+k, w_i] e^{jw_i n} \qquad (13)$$

Where, K= overlapping ratio (N/M).
N= no. of samples in each block
M= no. of sub bands
For windowing, used by hamming window, this has better selectivity for large signals [10].

IV. SUB BAND ADAPTIVE GSC

We can employ a GSC at every set of sub-bands for beam forming when we are capable to know the direction of arrival the signal and noise. The structure is shown in below Fig. 6.
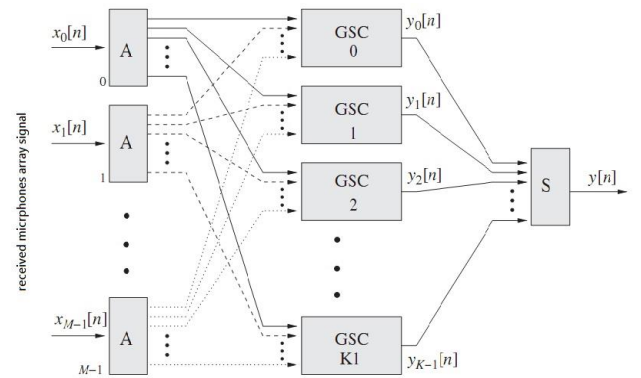

Fig.6. General SAB structure with a GSC in every sub-band [4].

In the above GSC, It was limited to using the same number of the analysis filter bank when the microphones number M will be same as well as the same number of GSC's as sub band number K. We have to divide each microphone signal into sub bands and implement a GSC to each of the responding sub bands. The computational load of the system will be high when the number of microphone arrays as well as sub band is high [4].

V. SYSTEM DESIGN
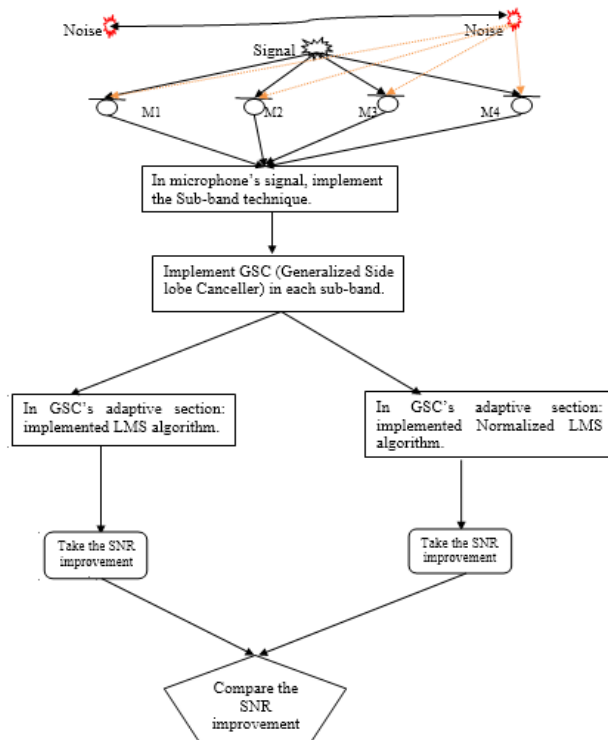
A. System Block Diagram

Fig.7. Flowchart of the overall system
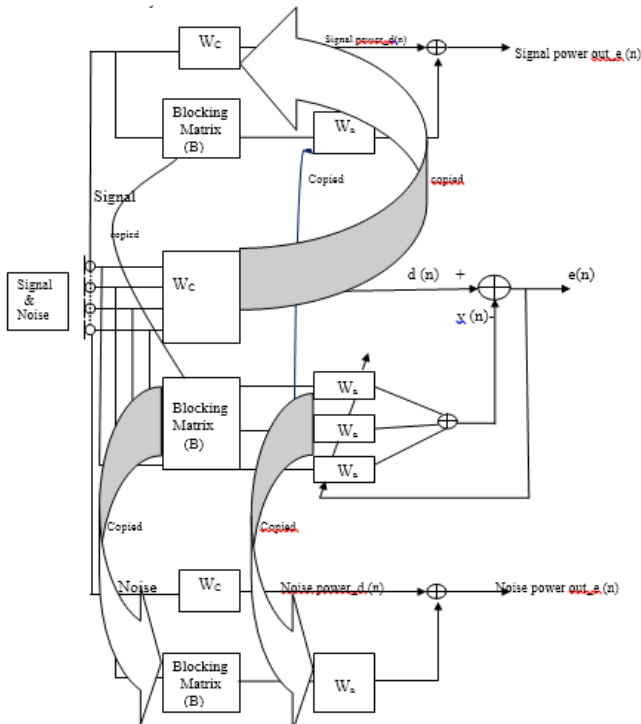
B. SNR Measurement Scenario

Fig.8. Overall SNR measurement setup

Based on the Fig.8, SNR calculations is like as following:

$$SNR1 = \frac{\text{Signal power d\_n}}{\text{Noise power d\_n}}$$

$$SNR2 = \frac{\text{Signal power out\_e (n)}}{\text{Noise power out\_e (n)}}$$

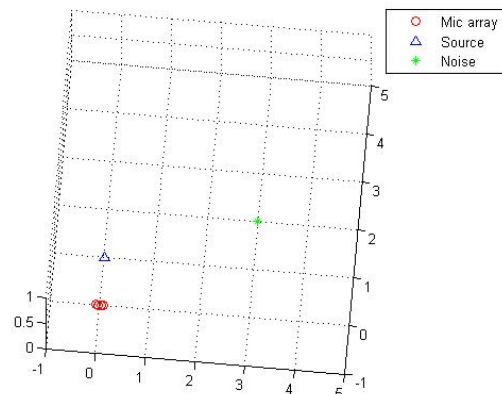## VI. RESULT CALCULATIONS AND COMPARISONS

Fig.9. Position of microphones, source and noise signal.

After implementing fractional time delay filtering in each of the microphones then signal was captured which is mixed up with noise [5].
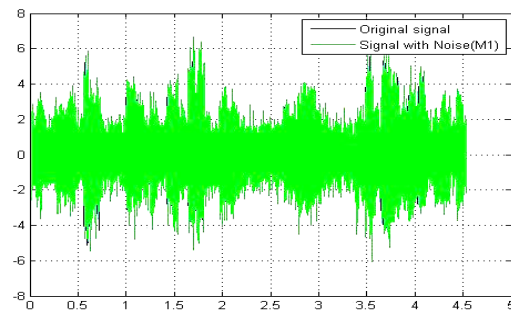

Fig.10. Original signal and signal with noise.

After receiving all the microphone signal with white noise, then we divided every sensor signal into sub bands and apply a GSC to each of the corresponding sub bands for the comparison of SNR improvement of speech signal. In the adaptive section of the SGC, the algorithm LMS and NLMS were implemented for noise reduction [5].
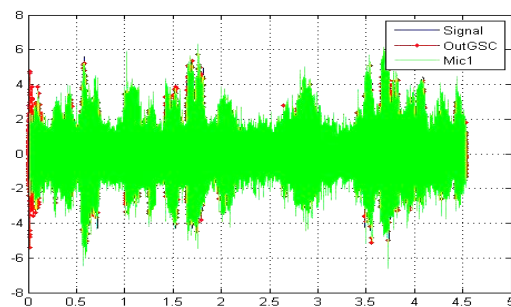

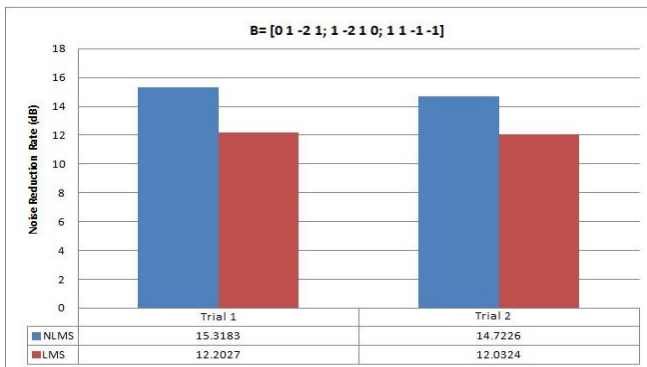Fig.11. Original signal, signal with noise and output at GSC.

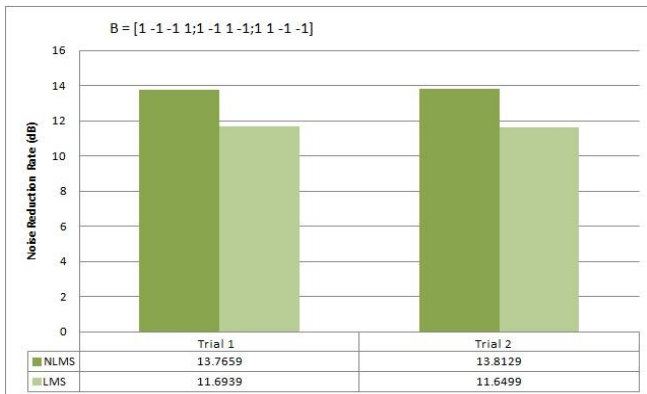Fig.12. Noise reduction comparison after using LMS and NLMS algorithm.

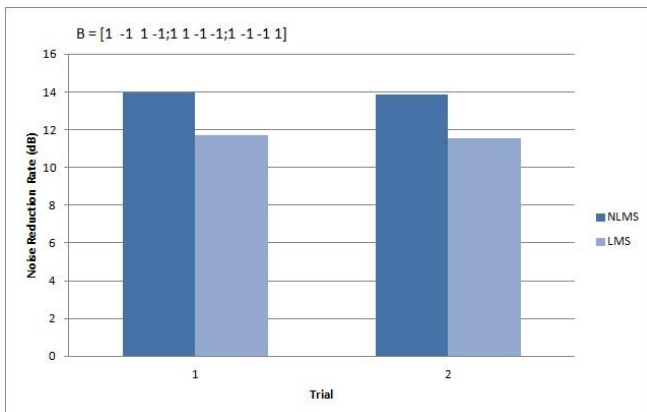Fig.13. Noise reduction comparison after using LMS and NLMS algorithm with the DFT matrix.

Fig.14. Noise reduction comparison after using LMS and NLMS algorithm with Hadamard matrix.

VII. Conclusion

In the above discussion, we have presented the analysis and comparison between LMS and NLMS algorithm for speech enhancement as well as noise suppression. To acquire the target we have used main signal mixed with noise signals. The overall signals were divided into sub bands and

transferred every sub bands into GSC. In the adaptive unit of GSC's, noise reductions were done by using LMS and NLMS algorithm. In the GSC's section, we have used different blocking matrix and SNR measurement has been taken on basis of the matrixes and comparative SNR result which was already shown on the graph. By using the NLMS and LMS algorithm, noise suppression levels up to 13 dB and about 10 dB. High SNR improvement has been shown while low SNR mixed up with the input signal.

References

[1] A. Abad, and J. Hernando. "Speech enhancement and recognition by integrating adaptive beamforming and Wiener filtering". In IEEE Sensor Array and Multichannel Signal Processing Workshop, SAM, Sitges. Citeseer, 2004.

[2] J.H. Lee, and C.L. Cho. "GSC-based adaptive beamforming with multiple-beam constraints under random array position errors". Signal processing, 84(2):341–350, 2004.

[3] M. Imani, and U.B.Neto, "Optimal gene regulatory network inference using the Boolean kalman filter and multiple model adaptive estimation." Signals, Systems and Computers, 2015 49th Asilomar Conference on. IEEE, pp. 423-427, 2015.

[4] W. Liu and S. Weiss. "Wideband Beamforming: Concepts and Techniques". Wiley, 2010.

[5] M. Ahmed, "Adaptive Sub band GSC Beam forming using Linear Microphone-Array for Noise Reduction/Speech Enhancement," M.S thesis, Dept. Signal Processing, Blekinge Institute of Technology, Sweden, 2012.

[6] M.H. Hayes. "Schaum's outline of theory and problems of digital signal processing". McGraw-Hill, 1999.

[7] L. Griffiths and CW Jim. "An alternative approach to linearly constrained adaptive beamforming". Antennas and Propagation, IEEE Transactions on, 30(1):27–34, 1982.

[8] P. Townsend, "Enhancements to the Generalized Side lobe Canceller for Audio Beam forming in an Immersive Environment," M.S thesis, University of Kentucky, Kentucky, UK, 2009.

[9] S. Hosseini, "Mapping Based Noise Reduction for Robust Speech Recognition," M.S thesis, Blekinge Institute of Technology, Blekinge, SE, July 2010.

[10] http://www.azimadli.com/Vibman/gloss_hammingwindow1.htm

Mamun Ahmed completed M.Sc in Signal Processing from BTH, Sweden and BSc in CSE from CUET, Bangladesh. He was a RF Engineer at Motorola, Bangladesh. Now, he is a faculty member of CSE Department at BAIUST, Bangladesh. His research interests are on adaptive signal processing techniques and wireless communication system/networks.

Nasimul Hyder Maruf Bhuyan completed MSc in Radio Communication from Blekinge Tekniska Högskola (BTH), Sweden. He completed his master's thesis from Lund University, Sweden. Now, he is working on sensor network and machine to machine communication. His research interests are on wireless sensor networks, network security and device to device communication.